

AVALIAÇÃO DE PRODUÇÃO ESCRITA: ASPECTOS A CONSIDERAR

Liv Rocha Fernandes (UNICAMP)

1. AVALIAR POR QUÊ?

Antes de iniciar um processo avaliativo, especialmente em língua estrangeira é necessário refletir. Por que avaliar, o que avaliar, como avaliar, como reportar as notas e dar uma respostas aos alunos, são importantes questões as quais o professor deve se ater antes de produzir ou aplicar uma prova. Mesmo que o processo de avaliação seja contínuo, e não tradicional, como colocam os PCNs (Brasil, 1999), que instrumentos utilizar para garantir uma avaliação válida e que produza respostas que auxiliem o professor, o aluno e a aula?

A avaliação faz parte da vida, ou “avaliar é um aspecto universal da vida social” (MCNAMARA, 2000, p.3). Segundo o autor, a avaliação permeia nossas vidas desde um simples exame de sangue, a um teste *antidoping* até os testes de DNA. Em outros casos, a avaliação tem um papel de *gatekeeper*, ou seja, uma triagem entre aqueles que desempenham adequadamente em certas situações ou funções, ou não. E de alguma forma esta visão, bem como o positivismo escolar, contribuem para que a avaliação seja temida, obscura e pouco compreendida.

A avaliação de produção oral em língua inglesa impõe ainda mais desafios. A LDB (1996) alega que o ensino de língua estrangeira é obrigatório e um direito de todo cidadão desde a 5ª série (agora 6º Ano do Ensino Fundamental). Entretanto, não há um currículo específico para língua inglesa, e sim os Parâmetros Curriculares Nacionais de 1999. Nos PCNs encontramos alegações do porquê a leitura é o foco do ensino de língua estrangeira moderna:

(...) Somente uma pequena parcela da população tem a oportunidade de usar línguas estrangeiras como instrumento de comunicação oral, dentro e fora do país. Mesmo nos grandes centros, o número de pessoas que utilizam o conhecimento das habilidades orais de uma língua estrangeira em situação de trabalho é relativamente pequeno. Deste modo, considerar o desenvolvimento de habilidades orais como central no ensino de Língua Estrangeira no Brasil não leva em conta o critério de relevância social para a sua aprendizagem. (...) O uso de língua estrangeira parece estar, em geral, mais vinculado à

leitura de literatura técnica ou de lazer. Nota-se também que os exames formais em Língua Estrangeira (vestibular e admissão a cursos de pós-graduação) requerem o domínio de habilidade de leitura. (P. 20)

A escrita dos PCNs não poderia prever que nosso país sediaria os Jogos Olímpicos (2014) e uma Copa do Mundo (2016), que se desenvolveria a ponto de sermos uma relevante força econômica no mundo e que a internet reconfiguraria as fronteiras globais, portanto a alegação de que os alunos usam ou usarão pouco as habilidade orais não vale mais. Mesmo assim, o foco é a leitura e o papel da oralidade fica em segundo plano, pois os PCNs argumentam que a oralidade teria a função de aumentar a consciência lingüística dos alunos e que momentos de oralidade “não implicam engajamento no discurso oral”. Desta forma, a oralidade é considerada importante, mas não essencial.

Uma questão essencial para a avaliação é a de que a produção oral (bem como a escrita) são avaliações de desempenho. Para McNamara (2000, p. 5), as avaliações diferem em método e propósito. Quanto ao método, as provas podem ser “lápiz e papel” ou de desempenho. As do primeiro tipo são usadas para avaliar “componentes separados de conhecimento de língua (gramática, vocabulário, etc.) ou de entendimento receptivo (compreensão auditiva ou de leitura” (ibidem). Já os testes de desempenho são de produção escrita ou oral e são avaliados em um ato de comunicação, onde o aluno dá uma amostra de discurso oral em uma tarefa que simula contextos do mundo real. Obviamente, não se avalia oralidade em provas de múltipla escolha ou de preencher lacunas.

Quanto ao propósito os testes podem ser de rendimento ou de proficiência. Os testes de rendimento estão ligados ao processo de instrução e podem ocorrer ao início, meio ou fim de um curso. Mais importante, ele deve “amparar o ensino ao qual se relaciona” (ibidem). Ao contrário, os testes de proficiência visam à situação futura de uso da língua (e.g.: vestibular ou pós-graduação).

Cabe dizer também que não é possível fazer uma prova perfeita. Assim o professor ou elaborador da prova oral deve privilegiar o aspecto que lhe parecer mais relevante. Refiro-me ao professor por ser ele muitas vezes o elaborador das aulas e do currículo em língua inglesa, conhecedor do contexto e dos alunos, por isso considero que ele seja a pessoa mais indicada a elaborar as provas de rendimento de seus alunos.

Logo, se cabe ao professor de língua desenvolver a prova e a partir daí, surgem outros pontos de interrogação. Qual o construto da prova? Que instrumentos usar? A nota será dada em número, em letras ou expressões (e.g.: “B+”; “8,0” ou “excelente”)? O que esperar do aluno e em que nota se encaixará a produção? Como será feito o *feedback*? Essas questões serão discutidas nos itens a seguir.

2. AVALIAR COMO?

Ao desenhar uma prova precisamos ter em mente qual é o construto dessa prova. O construto “se refere aos aspectos de conhecimento ou habilidade que o candidato possui e que são avaliados” (MCNAMARA, 2000, p. 13). Ou ainda segundo LUOMA (2005), o construto é o termo técnico que define o que se avalia, segundo a autora em avaliação de produção oral, se refere ao “tipo particular de oralidade”. Assim, no desenho da prova o professor deve refletir sobre que aspectos incluir no construto, ou seja, se ao avaliar produção oral, também se avalia gramática ou se nesta prova a gramática é irrelevante; se a interação faz parte da avaliação ou não, e se no momento da prova é mais importante ser assertivo quanto à formalidade ou à gramática ou ser comunicativo.

Em seu livro *Assessing Speaking* de 2005, a linguista aplicada e especialista em avaliação (*language testing*) Sari Luoma sugere um ciclo avaliativo, ilustrado pela figura 1. O ciclo se iniciaria quando surge a necessidade de avaliar, seguida por uma fase de planejamento e análise, que a autora chama de “*system development*”, na qual o construto é definido e onde tarefas, critérios classificatórios e procedimentos de administração são desenvolvidos, testados e revisados. Esses processos ajudam no monitoramento do ciclo e estabelecem procedimentos que garantem a qualidade. O próximo estágio é o de administração da prova, onde há interação do candidato/aluno com colega(s) ou com o examinador/professor. Então há a avaliação e classificação na qual o professor ou examinador usam o critério classificatório no desempenho mostrado pelo exemplo de discurso das tarefas da prova. Esse processo gera uma nota, que pode ser numérica ou uma resposta verbal, que deve satisfazer a necessidade inicial da prova, fechando assim o ciclo avaliativo.

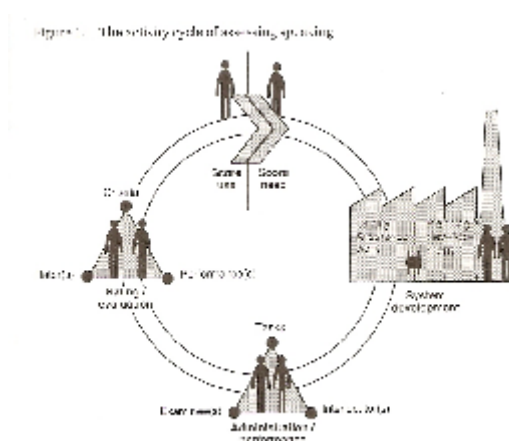


Figura 1.1- Ciclo avaliativo simplificado sugerido por LUOMA

No estágio de desenvolvimento a reflexão volta-se para o processo como um todo, como mostra a figura 1.2. O propósito do teste é pensado, e assim guia a criação das tarefas, critérios e instruções para o examinador e para os alunos e, como as notas serão relatadas e usadas.

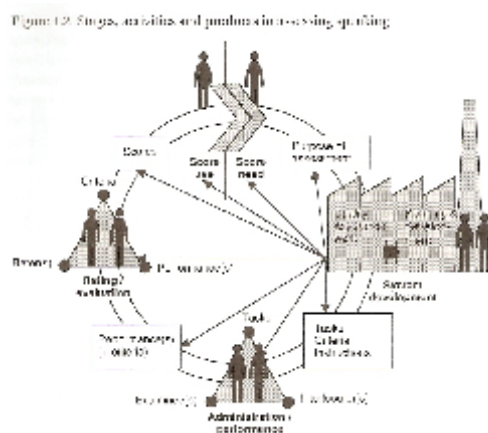


Figura 1.2 - Ciclo avaliativo simplificado sugerido por LUOMA

2.1. CRITÉRIOS

Os critérios se referem, segundo McNamara (2000, p.7) ao uso futuro “da vida real” da língua. Critério também pode ser entendido como “comportamento comunicativo relevante na situação-alvo” (ibidem) e ainda como um aspecto do

desempenho que é avaliado. Ainda segundo o autor, escolher os critérios envolve considerar que componentes de competência, como fluência, exatidão gramatical, sociocultural, formalidade, são importantes na prova oral. Balancear o peso de cada critério é uma discussão do desenvolvedor da prova ou do professor; escolher qual é relevância de cada um é um dos pontos de validação da prova. O construto sempre deve se relacionar diretamente e ser a base das escolhas de critérios.

A prova permite observar através de simulações e representações os critérios que antes não eram observáveis. Os desempenhos nas provas são usados para fazer inferências sobre o desempenho nos critérios, como podemos observar na figura 3, exemplo de McNamara (2000, p.8):

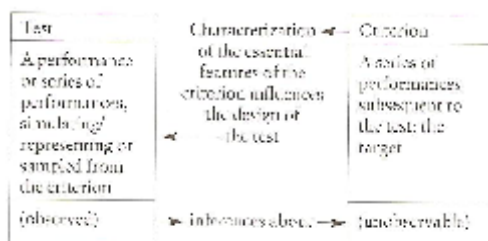


FIGURE 1.1 Test and criterion

Figura 3 – A relação entre critérios e prova

2.2. AS TAREFAS

Tarefas de produção, no ponto de vista de Luoma (2005, p. 31) são atividades que envolvem o falante com o propósito de atingir uma meta ou objetivo particular em uma situação de produção oral. Tal definição é orientada para o uso de língua alvo, e pode ser usado em situação de avaliação ou em outros momentos. Quando se desenha uma tarefa para avaliação é importante que as informações relacionadas ao construto apareçam através da nota. E isso, segundo a autora, dita o que tipos de habilidades a tarefa deve fazer o aluno mostrar.

A escolha da tarefa depende de que tipo de informação é preciso tirar das conclusões das notas e também do tipo de estrutura, podem ser: tarefas orais de resultado aberto ou orais estruturadas. As tarefas de resultado abertoⁱ “guiam as

discussões, mas permitem espaço para formas diferentes de cumprir os requisitos da tarefa (LUOMA, 2005, p. 48), em contrapartida, as tarefas estruturadas de produção oral delimitam o que o candidato deve dizer. A diferença também está na extensão do discurso do candidato, enquanto nas tarefas de resultado aberto espera-se um (ou vários) turno (s) longo de conversa; a tarefa estruturada de produção oral pede uma produção limitada.

As avaliações de produção oral têm tradição individual, ou seja, o candidato e um examinador, entretanto muitos críticos alegam que nesse tipo de arranjo o papel do candidato se resume a responder e assentir, o que não geraria evidência sobre suas habilidades comunicativas. Uma alternativa seria avaliar em pares ou em grupos. Em pares as vantagens seriam: a possibilidade de reunir mais evidências sobre as habilidades do candidato, bem como mais tipos de conversas; influenciar positivamente os alunos para o trabalho em pares nas aulas, ou mesmo representar o trabalho comum em pares de uma sala e, finalmente, economia de tempo dos examinadores.

A primeira desvantagem seria em relação ao examinador que não participa, apenas observa a “conversa” entre os pares, ação que artificializaria toda a interação. Outra preocupação é a de que nem todos os candidatos poderiam demonstrar suas habilidades orais da mesma forma, pois a fala do candidato é quase sempre influenciada pelo outro candidato (em termos de personalidade, estilo e nível lingüístico). Embora estudos tenham mostrado que a influência é pequena, uma sugestão de Luoma (2005, p.37) seria a de interpretar as notas de provas em pares com cuidado e se possível, fazer mais de um teste, no qual o candidato teria outro par. Outro problema seria o de que os candidatos podem não entender o funcionamento de uma entrevista, qual a responsabilidade de se falar em turnos, se podem interromper um ao outro, se podem fazer outro tipo de perguntar. Para tentar amenizar esse problema, quem faz a prova (ou o professor) deve se assegurar de que as instruções sejam dadas eficientemente, que os materiais da(s) tarefas(s) sejam claros e que os candidatos saibam que tipo de comportamento é esperado deles.

As interações em grupos não são muito comuns nas avaliações de produção oral, talvez por se difícil organizar o tamanho dos grupos e por misturar diferentes níveis de habilidade, capacidade, estilos de aprendizagem e personalidades. Entretanto, é uma

maneira eficiente, segundo Luoma (id., p.39) de amparar o ensino, por ser prático, por incentivar e gerar produção oral dos alunos.

2.3- AS ESCALAS DE CLASSIFICAÇÃO

O objetivo das escalas de classificação é orientar o processo de classificação das notas. As escalas de classificação, segundo Fulcherⁱⁱ (2003, p.88) que usa a definição de Davies, podem ser definidas como uma série de níveis construídos para julgar o desempenho dos candidatos. Esses níveis ou faixas vão de nenhum domínio a um ponto que representa um nativo educado.

Alguns críticos, como Lantolf e Frawley (apud FULCHER, 2003) afirmam que em escalas permeadas por este tipo de pensamento os alunos estão sendo avaliados por uma série de estágios discretos pela lei da não-contradição: alguém não pode estar em um nível e em outro, ou seja, um nível logicamente exclui todos os outros níveis. Além disso, as definições de nível são medidas por um “falante nativo educado” eles alegam que longe de ser um teste referenciado por critério, os conceitos de nível são de referências normativas. Também criticam a postura psicométrica, segundo eles a escala de classificação dá a ilusão de estar medindo, mas na verdade, a métrica não existe.

Há muita crítica envolvendo a questão de se avaliar um aprendiz de língua com um falante nativo, a primeira delas é o ponto da língua padrão (*standard*). Mesmo na língua nativa temos o problema de língua padrão, devido às inúmeras variedades e regionalidades. Desta forma, como determinar qual é a pronúncia padrão? Assim, Luoma (2005, p. 10) afirma que poucos seriam os alunos capazes de alcançar o padrão nativo em todos os aspectos o que implicaria na falha de muitos alunos se “falar como nativo” for um critério avaliativo. Assim, seria mais interessante avaliar a eficiência comunicativa, cuja base é a compreensão e definições de pronúncia em termos de rendimento realístico do aprendiz.

As escalas são, geralmente, usadas para classificar amostras de discurso de candidatos e guiar o professor ou o desenvolvedor da prova para a escolha das tarefas. As escalas podem ainda, orientar os usuários, seriam “usadas para relatar informações sobre comportamentos típicos do candidato em dado nível”; orientador os

examinadores, “desenhado para orientar o processo classificatório, focando na qualidade do desempenho esperado” e, orientador do desenvolvedor, que ajuda os construtores/desenvolvedores do exame a selecionar tarefas.

O nível de detalhamento de cada escala varia, sendo que do construtor/desenvolvedor é o mais completo e complexo. O nível para o usuário geralmente relata a nota em termos de faixa (ou nível) descritiva, que informa o que o candidato pode dizer na língua-alvo, já o nível de examinador contém a definição do construto, mas de forma concisa para que seja usada na classificação.

3. FEEDBACK

Outro passo importante é pensar no registro das notas, elas serão anotadas ou não e de que forma? Segundo Luoma (2005, p.172) o formulário ajuda na organização, na estruturação do processo e na consistência, além de “definir o que os avaliadores prestam atenção durante a avaliação. Formulários bem desenhados garantem que os aspectos pretendidos da fala são avaliados, e eles dão as informações de nota para relatar de forma prática.” Embora o professor ou o avaliador usem critérios e formulários, a nota informada é a geral, que de acordo com Luoma (id., p.173):

As notas gerais também servem como uma resposta (*feedback*) detalhada pode ser dada em termos de avaliação separada de aspectos analíticos como inteligibilidade, ritmo e entonação, exatidão gramatical, variação lexical e apropriabilidade da língua em uso, por exemplo. Se um formulário avaliativo foi usado, o retorno pode descrever os pontos fortes e fracos do desempenho verbal do candidato em mais detalhes.

Ainda para a autora, o *feedback* informal é muito comum nos contextos de ensino, apesar disso, o feedback “útil é concreto e descritivo”, ou seja, relaciona o desempenho do aluno às metas e descrições do que seria um bom desempenho. Dar exemplos ou descrições de bons desempenhos ajuda os alunos a definir metas de aprendizado e a entender melhor o processo de classificação.

Desta forma, comentar “bom” ou “ótimo” não daria ao aluno uma resposta satisfatória, pois não se pode entender o que se fez de certo ou errado, que comportamentos são reforçados ou devem ser repensados. Organizar boletins em termos

de metas de aprendizagem seria uma possibilidade para organizar melhor a resposta para os alunos, de acordo com Luoma (ibid., p. 189), mas o professor precisa definir quais são as metas para a produção oral em primeiro lugar.

Apesar do tomar tempo, detalhar o *feedback* pode ajudar alunos e professor no processo de ensino-aprendizado de língua. Um formulário de classificação bem feito pode amparar as sessões de *feedback*, que quando usadas com conselhos para o aprendizado a ser desenvolvido torna ainda mais eficiente o processo avaliativo.

4. CONCLUSÃO

A avaliação de produção oral realmente desafiador provoca grandes reflexões avaliativas e também sobre o conceito do que é língua e o uso de linguagem para o professor. Uma avaliação confiável deve levar em consideração estas questões e também outras. Avaliar não é dizer “está bom” ou “está ruim”. Significa entender porque se está avaliando, o quê se está avaliando e como se está avaliando. Muitas vezes nos deparamos com desempenhos orais de nossos alunos, que segundo nossos critérios não condizem com o que esperamos. Quase nunca paramos para refletir sobre nossa visão avaliativa, entretanto esta metarreflexão é necessária para compreendermos sob que pontos de vistas e critérios avaliamos cotidianamente nossos alunos e, principalmente, se nosso *feedback* é satisfatório, ou se estamos “dando nota” (me refiro a um tipo de classificação (*rating*) automática e sem critérios específicos).

É preciso um cuidado especial ao olhar a produção de nossos alunos. Muitas vezes, como educadores, em uma tarefa em que o foco é o sócio-cultural estamos avaliando exatidão gramatical, que é outro critério, que pode ou não estar no construto da prova. É preciso entender em que momento é importante avaliar o quê. Entretanto, se o professor entender que para se comunicar inteligivelmente o aluno precisa ter exatidão gramatical, o professor ou elaborador da prova deve retornar à etapa um envolver este aspecto no construto da prova. A validade da prova se dá pela qualidade do argumento que a embasa, e não é uma “verdade absoluta”. Por isso mesmo considero que as avaliações de produção oral em sala de aula deveriam ser desenhadas por quem mais conhece o contexto e os alunos, ou seja, pelo professor. Só o professor tem o

conhecimento e os argumentos para desenhar uma prova que seja adequada para cada turma, cada série, cada conjunto de alunos e para cada curso.

Espero que este pequeno artigo sirva de insumo para reflexões sobre a avaliação e que o professor ao ler, procure também mais informações, conceitos, questione e continue sempre aprendendo sobre a nossa profissão de ensinar.

REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. Parâmetros curriculares Nacionais - Terceiro e quarto ciclos do Ensino Fundamental – Língua Estrangeira, 1998.

BRASIL. Lei de Diretrizes e Bases da Educação, 1996.

FULCHER, Glenn. Testing Second Language Speaking. Pearson Longman, 2003.

LUOMA, Sari. Assessing Speaking. Cambridge University Press, 2005.

MCNAMARA, Tim. Language Testing. Oxford University Press, 2000.

NOTAS

ⁱ No original Luoma se refere a "open-ended speaking tasks", a tradução livre é minha.

ⁱⁱ Fulcher usa na página 88-89 uma citação de Davies et al. A referência para essa nota é: Davies et al. *Dictionary of Language Testing*. Cambridge University Press, 1999 (153-4).